



JP2000306030

Biblio Page 1

**DOCUMENT COLLATING DEVICE**

Patent Number: JP2000306030
Publication date: 2000-11-02
Inventor(s): KANEMOTO KOICHI; EGUCHI SHINICHI; OBARA KATSUTOSHI; CHIBA NOBUKAZU; YABUKI MASANORI; KATSUMATA YUTAKA
Applicant(s):: FUJITSU LTD
Requested Patent: JP2000306030 (JP00306030)
Application Number: JP19990118389 19990426
Priority Number (s):
IPC Classification: G06F19/00 ; G06K9/20
EC Classification:
Equivalents:

Abstract

PROBLEM TO BE SOLVED: To provide a document collating device for collating a registered document format with a read document, effective for a document to be identified unstably due to expansion/contraction of document, patchy lines or character contact.

SOLUTION: In this document collating device, a ruled line item including a reference point and rectangle coordinates in the document is registered, and a rectangle is extracted from the read document and collated while moving the reference point for each rectangle. Besides, the reference point is successively moved with the coordinates of the closest or adjacent rectangle as a reference. Further, the ruled line items of the adjacent rectangles are integrated and collated. In this case, while moving the reference point for each rectangle of the read document, concerning the rectangle of the registered document, ruled line information is collated for each rectangle in the order of rectangle numbers. A rectangle 1 is collated while using the origin of the document as the reference point.

Data supplied from the esp@cenet database - I2

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号
特開2000-306030
(P2000-306030A)

(43) 公開日 平成12年11月2日 (2000.11.2)

(51) Int.Cl. ⁷	識別記号	F I	テマコード*(参考)
G 0 6 F 19/00		G 0 6 F 15/22	3 1 0 Z 5 B 0 2 9
G 0 6 K 9/20	3 4 0	G 0 6 K 9/20	3 4 0 C

審査請求 未請求 請求項の数 4 O L (全 10 頁)

(21) 出願番号 特願平11-118389
 (22) 出願日 平成11年4月26日 (1999. 4. 26)

(71) 出願人 000005223
 富士通株式会社
 神奈川県川崎市中原区上小田中4丁目1番1号
 (72) 発明者 金元 浩一
 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内
 (72) 発明者 江口 真一
 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内
 (74) 代理人 100072590
 弁理士 井桁 貞一

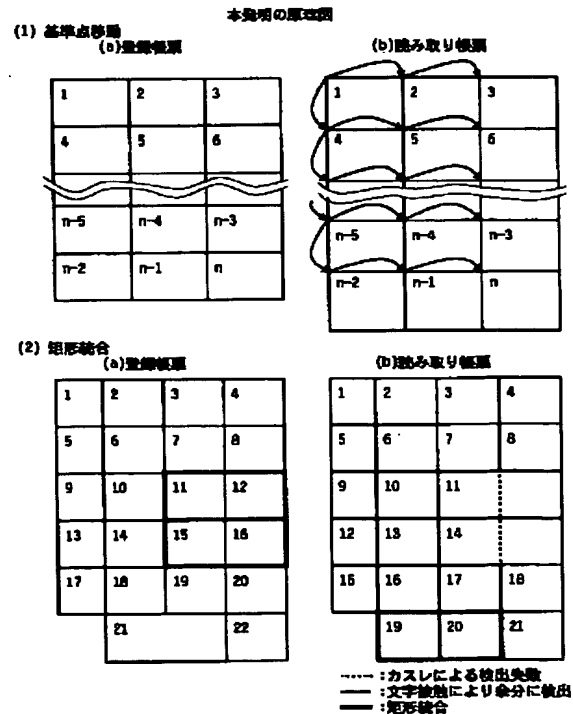
最終頁に続く

(54) 【発明の名称】 帳票照合装置

(57) 【要約】

【課題】 本発明は登録された帳票形式と読み取った帳票とを照合する帳票照合装置であって、帳票の伸縮や線のカスレや文字接触のために識別が不安定な帳票に対して有効な帳票照合装置を提供することを目的とする。

【解決手段】 本発明の帳票照合装置は、帳票内の基準点と矩形座標を含む罫線項目を登録しておき、読み取った帳票から矩形を抽出し、矩形ごとに基準点を移動させながら照合する。また最も近い矩形や隣接する矩形の座標を基準として基準点を順次移動させる。更に、隣接する矩形の罫線項目を統合して照合する。



【特許請求の範囲】

【請求項 1】 予め登録された帳票形式と読み取った帳票とを照合して帳票の種類を識別する帳票照合装置であつて、

帳票内の基準点と罫線で囲まれる矩形の座標とを含む罫線項目情報を帳票ごとに登録しておく記憶手段と、

読み取った帳票から罫線項目情報を抽出する抽出手段と、

前記抽出手段によって抽出された罫線項目情報と前記記憶手段に登録された罫線項目情報とを照合して一致した矩形の座標を次に照合する矩形の基準点として、基準点を順次移動させて照合する照合手段と、を備えることを特徴とする帳票照合装置。

【請求項 2】 罫線項目情報を照合して一致した矩形の内で最も近い位置にある矩形の座標を次に照合する矩形の基準点とすることを特徴とする請求項 1 に記載の帳票照合装置。

【請求項 3】 罫線項目情報を照合して一致した矩形の内で隣接する矩形の座標を次に照合する矩形の基準点とすることを特徴とする請求項 1 に記載の帳票照合装置。

【請求項 4】 予め登録された帳票形式と読み取った帳票とを照合して帳票の種類を識別する帳票照合装置であつて、

帳票内の基準点と罫線で囲まれる矩形の座標とを含む罫線項目情報を帳票ごとに登録しておく記憶手段と、読み取った帳票から罫線項目情報を抽出する抽出手段と、

読み取った帳票若しくは登録された帳票の内、いずれか一方の帳票に含まれる矩形の中から隣接する矩形の罫線項目情報を統合する統合手段と、

前記統合手段によって統合された罫線項目情報と他方の帳票の対応する罫線項目情報とを照合する照合手段と、を備えることを特徴とする帳票照合装置。

【発明の詳細な説明】**【0001】**

【発明の属する技術分野】 本発明は、複数種類の帳票がある場合に帳票処理を行うに先立って、読み取った帳票の種類を識別する技術に関わり、特に、帳票を読み取る時に帳票内の個々の矩形の識別が不安定な帳票に対して有効な帳票照合装置に関わる。

【0002】

【従来の技術】 図 8 の従来技術を用いて説明する。図 8 には処理対象の帳票イメージとデータベースが示される。帳票には帳票の基準点すなわち帳票原点と矩形の集合を含み、矩形は 2 本の縦罫線と 2 本の横罫線で形成される。矩形座標や罫線の属性などで表現される情報を罫線項目情報と呼ぶ。また、罫線項目情報の内、座標が一致する割合を罫線項目情報の一致率と呼ぶ。

【0003】 データベースには帳票の種類ごとの罫線項目情報が登録されている。帳票照合装置は処理対象の帳

票を読み取り、罫線項目を抽出すると共に、データベースから登録された罫線項目情報を読み出して罫線項目の照合処理を行う。

【0004】 従来の帳票照合装置では読み取った帳票を構成する矩形についての罫線項目情報を抽出し、データベースに既に登録済の帳票の種類ごとの罫線項目情報と照合し、個々の罫線項目情報の調べることによって帳票の識別を行っている。

【0005】 従来の罫線項目情報を利用した帳票照合では、読み取った帳票の傾きを補正すると共に帳票における原点位置の補正を行い、登録された帳票と読み取った帳票との罫線項目情報に基づいて座標照合を行っていた。帳票照合装置は罫線項目を照合した結果、一致した帳票情報若しくは一致率の高い帳票を出力する。

【0006】

【発明が解決しようとする課題】 しかしながら、帳票上の 1 カ所を基準点（帳票原点）にして各矩形の罫線項目の補正を行い帳票の照合を行う場合には、読み取り帳票上の矩形座標は帳票原点から離れる程誤差が大きくなり、正確な比較が難しくなるという問題がある。

【0007】 その上、帳票の読み取り時に帳票の伸縮がある場合は罫線項目情報の一致率が低くなる。また、矩形座標を照合するときのしきい値を緩めると帳票原点から遠い位置にある矩形の罫線項目の一致が取れてしまい、異なる帳票を一致していると見なしてしまうという問題がある。

【0008】 図 9 に伸縮により問題となる帳票例を示す。図 9 の (a) はデータベースに登録された登録帳票であり、正しい罫線項目情報がデータベースに登録されている。図 9 の (b) は読み取りされた照合帳票であり、矩形の座標が登録帳票に比べて帳票原点から縦方向に伸びている。罫線項目情報は帳票原点に近いと登録帳票とのズレは少ないが、帳票原点から離れるほど登録帳票とのズレが大きくなることを示す。

【0009】 また、帳票を読み取る際に罫線がカスレていると矩形の一部線分が抽出されずに、実際より大きな矩形と認識されてしまうという問題がある。

【0010】 更に、帳票に記述された文字が罫線と接触している場合には矩形の一部線分として抽出されてしまい、本来は一つの矩形である筈のところが小さな矩形が複数あるように認識されてしまうという問題がある。

【0011】 図 10 にカスレと文字接触により問題となる帳票例を示す。図 10 の (a) はデータベースに登録された登録帳票であり、正しい罫線項目情報がデータベースに登録されている。図 10 の (b) は読み取りされた帳票であり、カスレと文字接触により安定しない罫線項目情報を示す。

【0012】 読み取り帳票の矩形 11 と矩形 14 にはカスレにより検出できなかった罫線項目を点線で示してある。また、読み取り帳票の矩形 19 と矩形 20 の間には

文字接触などにより余分に抽出した罫線項目を二重線で示してある。

【0013】帳票のカスレにより罫線項目を正しく取得できなかった場合や帳票に記述された文字が罫線項目情報を複数に分離した場合は、読み取った帳票の登録された帳票に対する相違度が上がって正しく照合を行うことができない。

【0014】このように罫線項目情報による座標照合が安定しない場合は、その結果として帳票を構成する矩形の識別が不安定になる。

【0015】本発明はこのように点にかんがみて、帳票を読み取る際に帳票の伸縮や罫線のカスレがあったり、帳票に記述された文字が罫線に接触したりしているなどのために、帳票内の個々の矩形の識別が不安定な帳票に対して有効な帳票照合装置を提供することを目的とする。

【0016】

【課題を解決するための手段】本発明の帳票照合装置は、予め登録された帳票形式と読み取った帳票とを照合して帳票の種類を識別する帳票照合装置であって、帳票内の基準点と罫線で囲まれる矩形の座標とを含む罫線項目情報を帳票ごとに登録しておく記憶手段と、読み取った帳票から罫線項目情報を抽出する抽出手段と、抽出手段によって抽出された罫線項目情報と記憶手段に登録された罫線項目情報とを照合して一致した矩形の座標を次に照合する矩形の基準点として、基準点を順次移動させて照合する照合手段と、を備えることを特徴とする。

【0017】また、本発明の帳票照合装置は、罫線項目情報を照合して一致した矩形の内でも近い位置にある矩形の座標を次に照合する矩形の基準点とすることを特徴とする。

【0018】更に、本発明の帳票照合装置は、罫線項目情報を照合して一致した矩形の内でも隣接する矩形の座標を次に照合する矩形の基準点とすることを特徴とする。

【0019】更にまた、本発明の帳票照合装置は、予め登録された帳票形式と読み取った帳票とを照合して帳票の種類を識別する帳票照合装置であって、帳票内の基準点と罫線で囲まれる矩形の座標とを含む罫線項目情報を帳票ごとに登録しておく記憶手段と、読み取った帳票から罫線項目情報を抽出する抽出手段と、読み取った帳票若しくは登録された帳票の内、いずれか一方の帳票に含まれる矩形の中から隣接する矩形の罫線項目情報を統合する統合手段と、統合手段によって統合された罫線項目情報と他方の帳票の対応する罫線項目情報とを照合する照合手段と、を備えることを特徴とする。

【0020】図1は本発明の原理図である。図1の

(1)は本発明の帳票照合装置の基準点移動を説明するための図である。図1の(1)の(a)は登録された帳票であり、合計n個の矩形が帳票を構成しており、横方向に例えば3個の矩形が並んでいる。矩形には1からn

までの矩形番号を付けた。図1の(1)の(b)は読み取った帳票であり、縦方向に矩形の位置がずれて並んでいる。

【0021】本発明の帳票照合装置では、読み取った帳票の矩形ごとに基準点を移動させながら矩形ごとに登録帳票の矩形について、矩形番号の順に罫線項目情報を照合する。矩形1の照合は帳票の原点を基準点にして行う。

【0022】図1の(1)の(b)に矢印で示したように、矩形2と矩形3は帳票に含まれる矩形の中から隣接する矩形、すなわち矩形1と矩形2を抽出し、それぞれの矩形の座標を基準点として基準点を順次移動させる。

【0023】矩形3から矩形4に移るときは、帳票に含まれる矩形の中から最も近い位置にある矩形の罫線項目情報を抽出する。この場合は矩形3よりも矩形1の方が矩形4に近いので矩形1の座標を用いて矩形4の照合を行う。

【0024】このように帳票上の原点とは別に矩形照合用の原点(矩形の基準点)を使用することにより、帳票上の1箇所の帳票原点を基準点にするよりも、矩形の各座標と基準点との距離を小さくできる。

【0025】これにより、読み取り誤差を少なくすることができるので、読み取った罫線画像(イメージ)に伸縮があっても正しく照合することが可能になる。

【0026】図1の(2)は本発明の帳票照合装置の矩形統合を説明するための図である。図1の(2)の

(a)は登録された帳票の例であり、大中小の3種類の矩形が合計22個の矩形によって帳票を構成している。図1の(2)の(b)は読み取った帳票の例である。

【0027】読み取った帳票は矩形番号11と14の点線で示した縦罫線がカスレのために検出できなかったことを示している。また矩形番号19と20の間の二重線は文字接触のために縦罫線を余分に誤って検出してしまったことを示している。

【0028】矩形番号11では読み取り帳票と登録帳票との罫線項目を照合しても一致しないので、小さい方の矩形(この場合は登録帳票の矩形11)を隣接する矩形(登録帳票の矩形12)と統合した罫線項目と、読み取り帳票の矩形11の罫線項目とを照合する。

【0029】このように照合した矩形の内、小さい方の矩形を隣接する矩形と統合して照合すれば一致するので、大きい方の矩形11(読み取り帳票)の縦罫線が検出できなかったことが判明する。

【0030】同様に、登録帳票の矩形15でも一致しないので、登録帳票(小さい方)の矩形15と隣接する矩形16とを統合し、読み取り帳票の矩形14の罫線項目とを照合する。この場合は、読み取り帳票の矩形14の縦罫線が検出できなかったことが判明する。

【0031】また、登録帳票の矩形番号21では読み取り帳票の矩形番号19と罫線項目を照合しても一致しな

いので、小さい方の矩形（この場合は読み取り帳票の矩形19）を隣接する矩形（読み取り帳票の矩形20）と統合した罫線項目を登録帳票の矩形21の罫線項目と照合する。この場合も小さい方の矩形を統合して照合すれば一致するので、読み取り帳票の矩形19と矩形20の間に余分な縦罫線を検出してしまったことが判明する。

【0032】このように登録帳票と読み取り帳票を照合して罫線項目が一致しなかった場合は、小さい方の矩形を隣接する矩形と統合することにより、大きい方の矩形の縦罫線が読み取りに失敗したのか、または、余分に検出してしまったのかを判別できる。

【0033】これにより、読み取った帳票に罫線のカスレや文字接触があつて、罫線項目が正しく検出できなかった場合でも安定して帳票を照合することが可能になる。

【0034】図2の本発明の帳票照合装置の構成図を用いて本発明の帳票照合装置の構成を説明する。図2の1は帳票照合装置、2は記憶手段、3は帳票読み取り手段、4は抽出手段、5は統合手段、6は照合手段である。

【0035】帳票照合装置1は帳票読み取り手段3により帳票を読み取り、記憶手段2に予め登録された帳票と照合する。記憶手段2には帳票を構成する矩形についての罫線項目情報を帳票ごとに登録しておく。

【0036】罫線項目情報には罫線の種類などの属性と帳票内の2本の縦罫線と2本の横罫線で囲まれる矩形の座標と帳票内の基準点とを含む。

【0037】抽出手段4は読み取った帳票から矩形ごとの罫線項目情報を抽出する。照合手段6は、抽出手段4によって抽出された罫線項目情報と記憶手段2に登録された罫線項目情報とを矩形ごとに基準点を移動させながら矩形ごとに罫線項目情報を照合する。

【0038】抽出手段4は帳票に含まれる矩形の中から隣接する矩形の罫線項目情報を抽出し、照合手段6は隣接する矩形の座標を基準点として基準点を順次移動させて照合する。

【0039】また、帳票内の矩形を照合して一致しない場合は、統合手段5は帳票読み取り手段3で読み取った帳票、若しくは記憶手段2に登録された帳票の内、いずれか一方の帳票に含まれる小さい方の矩形の中から隣接する矩形の罫線項目情報を統合する。

【0040】この場合、照合手段6は統合手段5によって統合された罫線項目情報と他方の帳票の対応する大きい方の矩形の罫線項目情報とを照合する。

【0041】帳票の形式と本発明の基準点について図3の帳票形式と原点移動を用いて説明する。帳票には帳票原点があり、帳票の形式は図3のようにいくつかの矩形群から構成されることが多い。図3では帳票は3つの矩形群(A)、(B)、(C)から成り、点Gが帳票原点を示す。

【0042】帳票内の最初の矩形群(A)の先頭の矩形a1を照合するときは帳票原点Gを基準にする。矩形a1の罫線項目を照合して一致すれば矩形群(A)の次の矩形は照合して一致した矩形の内でも近い位置にある矩形、すなわち矩形a1の座標を基準点とする。矩形を照合して次々に一致すれば、矩形群の中で隣接する矩形の座標が次に照合する矩形の基準点になる。

【0043】矩形群から次の矩形群に移るときも、それまでに照合して一致した矩形の内でも近い位置にある矩形の座標を基準点とする。従って矩形群(A)の矩形a2が照合の結果、一致していれば矩形群(B)の先頭の矩形b1は矩形a2の座標を基準にする。矩形群(B)の矩形b2が照合の結果、一致していれば矩形群(C)の先頭の矩形cは矩形b2の座標を基準にする。

【0044】

【発明の実施の形態】本発明の実施例について図面を用いて詳細に説明する。図4は原点移動による処理フロー図である。第1の発明の実施例について図4の処理フロー図を用いて説明する。まず、データベースに登録された帳票と読み取った帳票について最初の矩形を照合する。(ステップS410)

矩形番号をKとして $K=1$ とし、原点移動のときに用いる原点番号をGとして $G=0$ とする。最初の矩形座標は帳票原点($G=0$)を基準点にして登録された帳票と読み取った帳票の罫線項目情報を比較する。(ステップS410)

最初の矩形の照合の結果、罫線項目情報が一致していればステップS430に進み、一致していなければステップS425に進む。

【0045】帳票を照合した最初の矩形が不一致だった場合はデータベースから他の登録帳票を読み出し、改めて照合処理を開始する。(ステップS425からステップS410に戻る。)

帳票の最初の矩形が一致した場合は原点番号Gを1つ更新して $G=1$ とする。原点番号を1にしたのは、2つ目の矩形の基準点を最初の矩形($K=1$)の座標にするためである。(ステップS430)

続いて登録帳票の矩形番号が最後であるか否か、すなわち $K=n$ であるか否かを判定する。ステップS440で最後の矩形であれば照合処理を終了する。最後の矩形でなければステップS450に進む。

【0046】ステップS450では次の矩形の座標を照合するために矩形番号Kを1つ更新して $K=K+1$ とする。次に $K(=K+1)$ 番目の矩形とG番目の矩形がそれまでに照合して一致した矩形の中で最も近い位置にあるか否かを判定する。(ステップS460)

帳票に多数の矩形がある場合は矩形番号が連続していても $K+1$ 番目の矩形の基準点が常にK番目の矩形座標になるとは限らない。矩形番号を更新するときに矩形群をまたがったり、同じ矩形群内であっても行や列が改めら

れることがある。また、K番目の矩形が一致しないので基準点にできない場合もある。

【0047】そこでステップS465では原点番号Gをこれから照合する矩形(矩形番号K+1)に最も近い位置にある矩形のものとするように更新する。それまでに照合して一致した矩形の中で行や列が同じものがあれば、その矩形を選択する。例えば行が変わった場合は列が同じもの(一つ前の行の矩形であることが多い。)があれば、その矩形番号を原点番号Gにセットする。

【0048】ステップS470では原点番号Gの矩形座標を基準点として、登録帳票のK(=K+1)番目の矩形と読み取った帳票の矩形との罫線項目を比較する。

【0049】登録帳票のK+1番目の矩形と読み取った帳票の矩形と照合した結果、罫線項目情報が一致した場合は一致したことを記録しておき、原点番号Gを更新してG=K(K=K+1)とする。(ステップS490)登録帳票のK+1番目の矩形と読み取った帳票の矩形と照合した結果、罫線項目情報が一致していない場合は不一致であると判定したことを記録しておく。この場合、原点番号Gは更新しないで次の矩形に進む。(ステップS495)

このように罫線項目を照合して不一致が発生した場合は不一致であった矩形をスキップするので、原点移動による照合処理が終了したときには矩形番号Kと原点番号Gの差に基づいてその帳票における罫線項目情報の一致率が判明する。

【0050】次に矩形番号と本発明の基準点について図5の矩形番号と基準点を用いて説明する。矩形番号の採り方は図5の(a)、(b)、(c)などがある。図5の(a)では、1行目の矩形に左から右に番号を付与しており、各矩形の基準点は左から右に移動する。2行目の先頭の矩形5は矩形1の座標を基準点にする。

【0051】以下同様にして順次基準点を移動し、基準点にする矩形座標と次に照合する矩形座標が同じ行にある場合も、改行された場合も隣接する矩形の座標を基準点にして次の矩形の罫線項目を比較する。

【0052】図5の(b)では、1列目の矩形に上から下に番号を付与しており、各矩形の基準点は上から下に移動する。2列目の先頭の矩形5は矩形1の座標を基準点にする。

【0053】以下同様にして順次基準点を移動し、基準点にする矩形座標と次に照合する矩形座標が同じ列にある場合も、列を改めた場合も隣接する矩形の座標を基準点にして罫線項目を比較する。

【0054】図5の(c)では、1行目の矩形に左から右に番号を付与しており、各矩形の基準点は左から右に移動する。2行目の先頭の矩形5は隣接する矩形4の座標を基準点にする。2行目の矩形には右から左に番号を付与しており、各矩形の基準点は右から左に移動する。

【0055】以下同様にして順次基準点を移動し、基準

点にする矩形座標と次に照合する矩形座標が同じ行にある場合も、行を改める場合も隣接する矩形の座標を基準点にして罫線項目を比較する。

【0056】図6は矩形統合による処理フロー図である。第2の発明の実施例について図6の処理フロー図を用いて説明する。まず、データベースに登録された罫線帳票と読み取った罫線帳票について、最初の矩形を照合する。(ステップS610)矩形番号をKとしてK=1とする。また、最初の矩形座標は帳票原点を基準点にして登録された罫線帳票と読み取った罫線帳票の罫線項目情報を比較する。

【0057】照合の結果、罫線項目情報が一致していればステップS700に進み、一致していなければステップS630に進む。(ステップS620)

比較の結果、一致しなかった場合は矩形の大小を比較し、小さい方の矩形とその矩形と隣接する矩形を統合する。次に統合した矩形の罫線項目と、大小を比較した大きい方の矩形の罫線項目と照合する。(ステップS630)

大きい方の矩形がデータベースに登録された帳票のものであるとは限らない。

【0058】照合の結果、罫線項目情報が一致していればステップS650に進み、一致していなければステップS660に進む。(ステップS640)

統合した方の矩形が登録された帳票である場合は、読み取り帳票の矩形の縦罫線がカスレなどのために読み取りできなかったことが推定される。この場合は登録帳票のK番目の矩形が読み取り帳票の矩形より小さい矩形であったので、K番目とK+1番目の矩形を統合したことになる。従って矩形番号Kを1つ更新してK=K+1とする。(ステップS650)

大小を比較したときの大きい方の矩形が登録された帳票である場合は、読み取り帳票の矩形部分に文字接触などがあり、そのために矩形を正しく読み取りできなかったことが推定される。この場合は登録帳票のK番目の矩形が読み取り帳票の矩形より大きい矩形であったので、読み取り帳票の矩形を統合したことになる。従って矩形番号Kは更新しなでK=K+0とする。(ステップS650)

1回目の統合をして照合した結果、罫線項目情報が一致していない場合は、矩形の大小を比較したときの、大きい方の矩形とその矩形と隣接する矩形を統合する。(2回目の統合)次にその統合した矩形の罫線項目と、ステップS630で統合した矩形の罫線項目とを照合する。(ステップS660)

2回目の統合を行い、照合の結果、罫線項目情報が一致していればステップS680に進み、一致していなければステップS690に進む。

【0059】2回統合を行い、登録帳票も読み取り帳票も統合して罫線項目が一致した場合は、縦罫線のカスレ

と文字接触が同時に発生したことが推定される。この場合、登録帳票のK番目とK+1番目の矩形を統合したことになり、K+1番目まで照合できたことになる。従って矩形番号Kを1つ更新して $K=K+1$ とする。(ステップS680)

登録帳票のK番目とK+1番目の矩形を統合し、読み取り帳票も統合したが、罫線項目が一致しない場合は、K+1番目に対応する読み取り帳票の矩形にカスレや文字接触が更に発生していることが推定される。この場合、K+1番目まで照合したが不一致であったと判定し、次のK+2番目以降の矩形を照合する。従って矩形番号Kを1つ更新して $K=K+1$ とする。(ステップS690)

続いて矩形番号が最後であるか否か、すなわち $K=n$ であるか否かを判定して最後の矩形であれば照合処理を終了し、最後の矩形でなければステップS710に進む。

(ステップS700)

矩形番号Kを1つ更新して $K=K+1$ とする。ステップS710でKを1つ更新することで次の矩形の照合処理をすることができる。ステップS650、ステップS680、またはステップS690で $K=K+1$ としていた場合は、次はK+2番目の矩形から照合処理を行うことができる。ステップS710からステップS620に戻る。

【0060】次に読み取り帳票の罫線のカスレと文字接触について図7の矩形統合と障害モードを用いて説明する。読み取り帳票の罫線のカスレと文字接触などの障害のモードの例を図7の(a)、(b)、(c)、(d)、(e)に示す。

【0061】図7のK、K+1は登録帳票の矩形番号、Y、Y+1は読み取り帳票の矩形番号、点線で示した縦罫線はカスレのために検出できなかったことを示し、二重線は文字接触のために縦罫線を余分に誤って検出してしまったことを示す。

【0062】図7の(a)は読み取り帳票の矩形Yがカスレのために登録帳票の矩形Kよりも大きくなっている。この場合は矩形Kと矩形K+1を統合して矩形Yと照合する。図7の(b)は読み取り帳票の矩形Yが文字接触のために登録帳票の矩形Kよりも小さくなっている。この場合は矩形Yと矩形Y+1を統合して矩形Kと照合する。

【0063】図7の(c)は読み取り帳票の矩形Yがカスレのために登録帳票の矩形Kよりも大きくなってお

り、更に文字接触のために登録帳票の矩形K+1よりも小さい矩形Y+1になっている。この場合は矩形Kと矩形K+1を統合したものと矩形Yと矩形Y+1とを統合したものを照合する。

【0064】図7の(d)は読み取り帳票の矩形Yが文字接触のために登録帳票の矩形Kよりも小さくなっており、矩形Y+1がカスレのために登録帳票の矩形K+1よりも大きくなっている。この場合は矩形Kと矩形K+1を統合したものと矩形Yと矩形Y+1とを統合したものを照合する。

【0065】図7の(e)は読み取り帳票の矩形Y+1が2本の縦罫線のカスレのために登録帳票の矩形K+1よりも大きくなっている。この場合は矩形Kと矩形K+1を統合したものと矩形Yと矩形Y+1とを統合したものを照合しても一致しない。図7の(a)、(b)、(c)、(d)、(e)の障害モードを検出するルーチンを図6に示す。

【0066】

【発明の効果】以上の説明から明らかなように本発明によれば、帳票を読み取る際に帳票内の矩形の識別が不安定になるという要因が発生しても、帳票の伸縮に対しては基準点の移動により、また、線カスレや文字接触に対しては矩形統合により、罫線項目の照合を行うことができ、安定的に帳票を識別できるという効果がある。

【図面の簡単な説明】

【図1】 本発明の原理図

【図2】 帳票照合装置の構成図

【図3】 帳票形式と原点移動

【図4】 原点移動の処理フロー図

【図5】 矩形番号と基準点

【図6】 矩形統合の処理フロー図

【図7】 矩形統合と障害モード

【図8】 従来技術

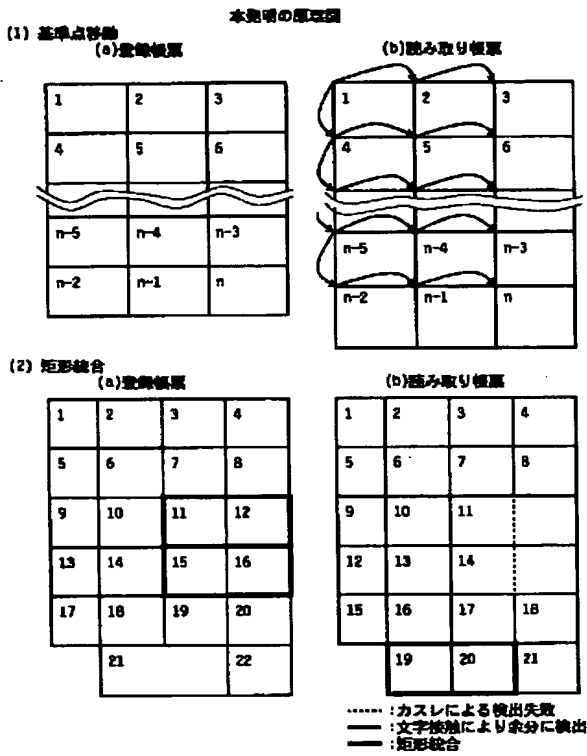
【図9】 伸縮により問題となる帳票例

【図10】 カスレと文字接触により問題となる帳票例

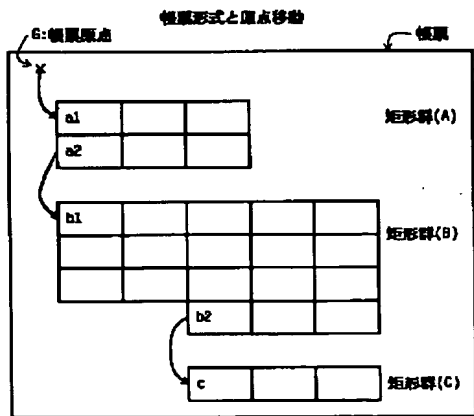
【符号の説明】

- 1 帳票照合装置
- 2 記憶手段
- 3 帳票読み取り手段
- 4 抽出手段
- 5 統合手段
- 6 照合手段

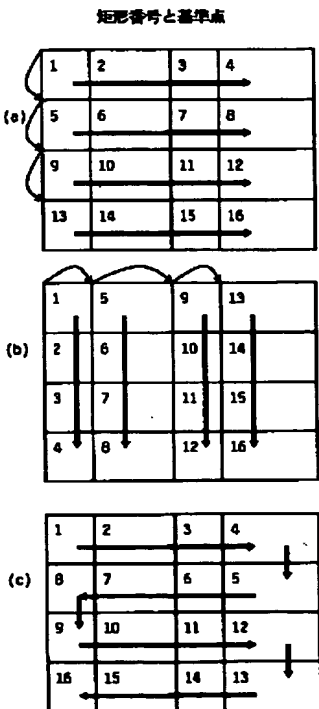
【図 1】



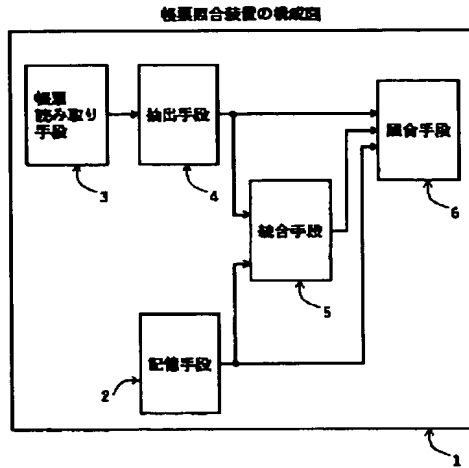
【図 3】



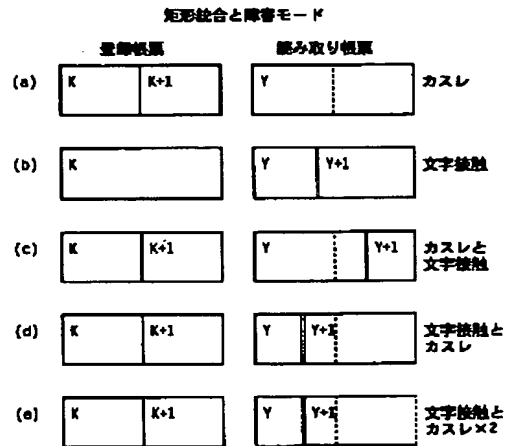
【図 5】



【図 2】

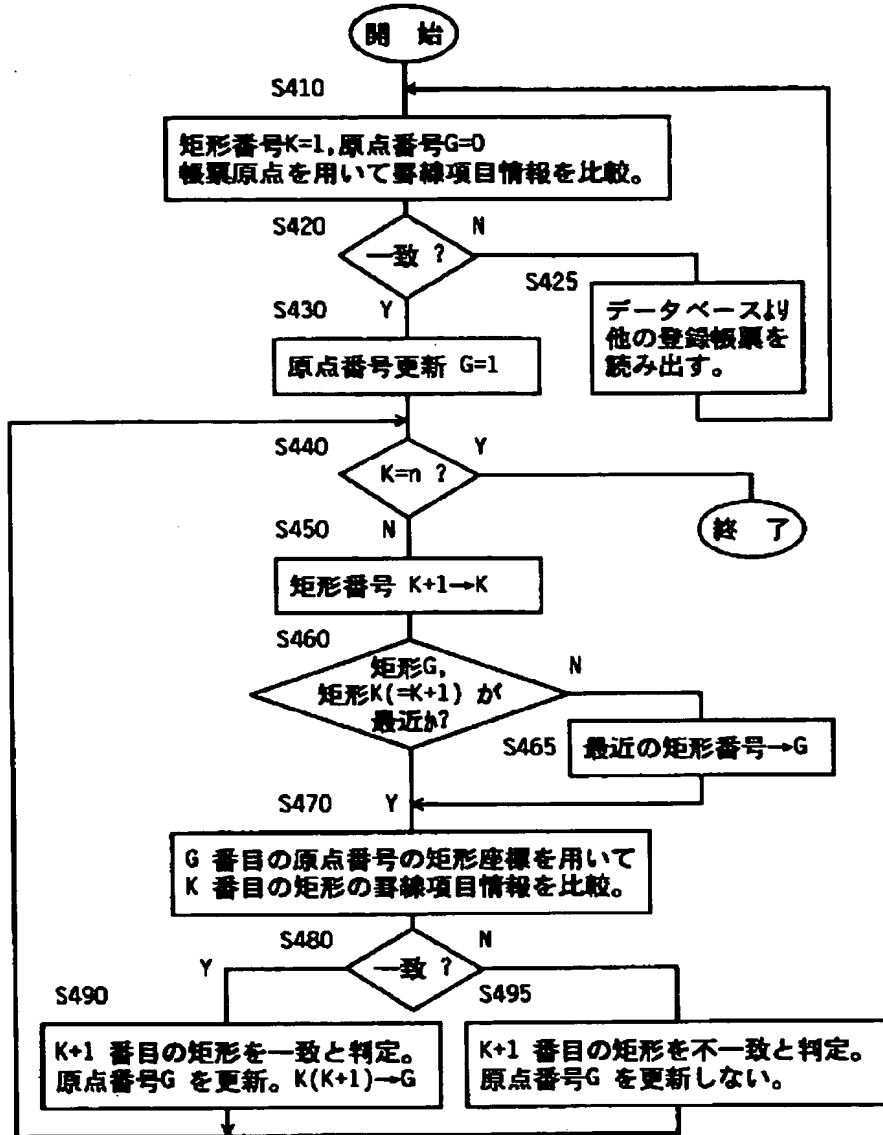


【図 7】



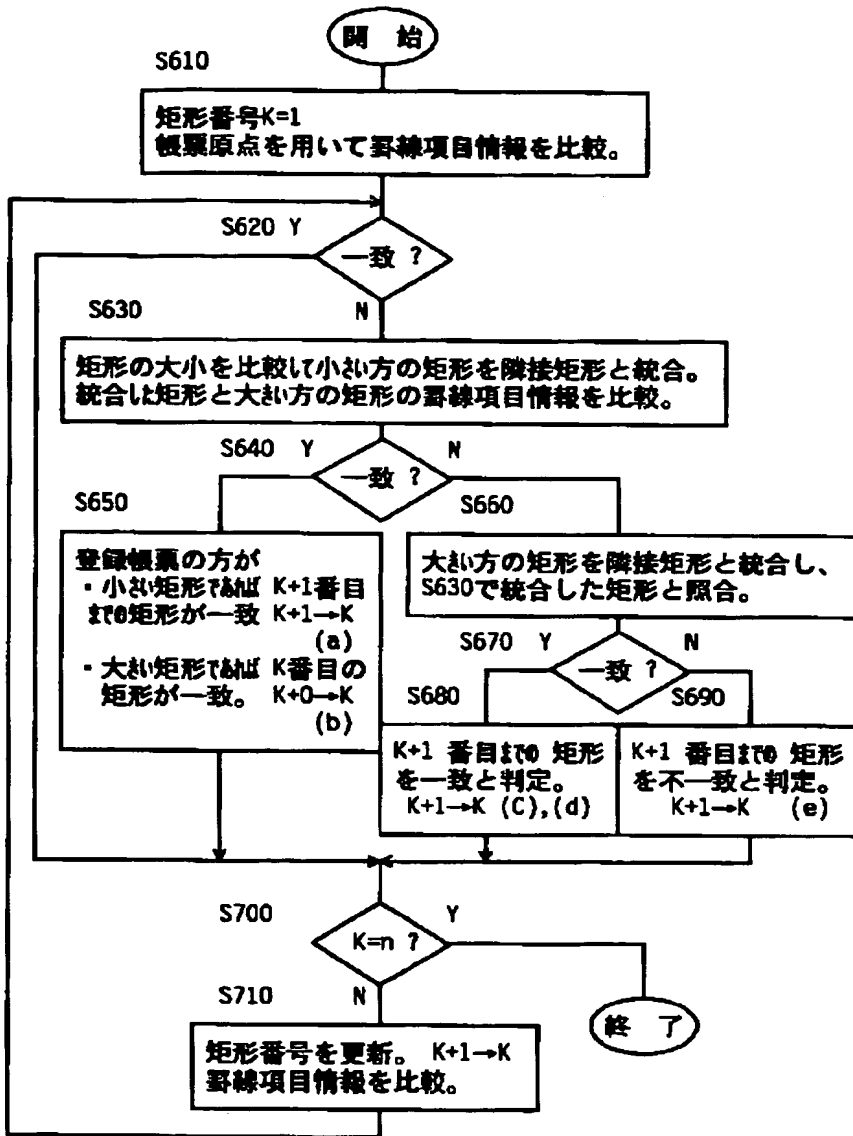
【図4】

原点移動の処理フロー図

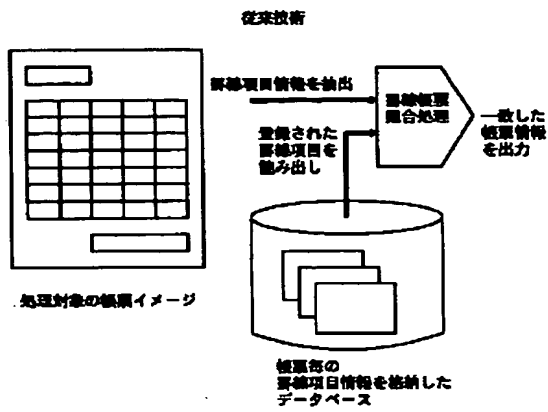


【図 6】

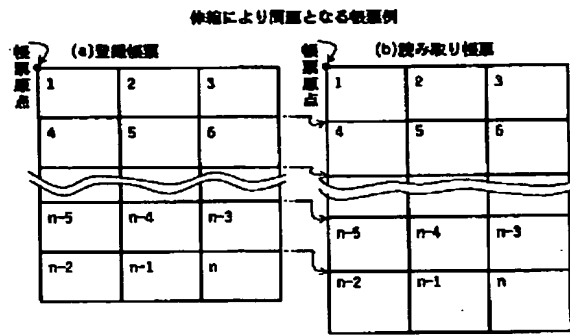
矩形統合の処理フロー図



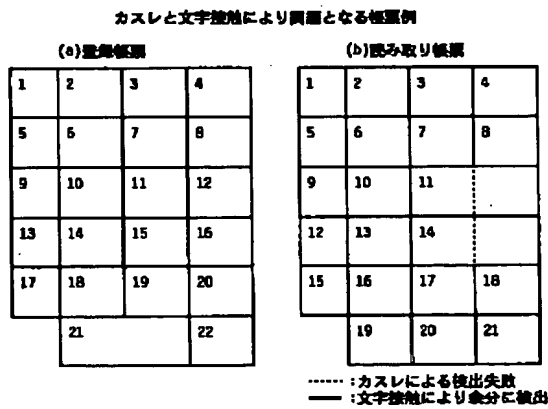
【図8】



【図9】



【図10】



フロントページの続き

(72) 発明者 小原 勝利
 神奈川県川崎市中原区上小田中4丁目1番
 1号 富士通株式会社内

(72) 発明者 千葉 亘一
 神奈川県川崎市中原区上小田中4丁目1番
 1号 富士通株式会社内

(72) 発明者 矢吹 真紀
 神奈川県川崎市中原区上小田中4丁目1番
 1号 富士通株式会社内

(72) 発明者 勝又 裕
 神奈川県川崎市中原区上小田中4丁目1番
 1号 富士通株式会社内

Fターム(参考) 5B029 BB02 CC18 CC27 CC30 EE12